Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nonlinear Analysis 63 (2005) e2057 – e2065
Nonlinear Analysis
www.elsevier.com/locate/na

A separation principle for partially observed control of singular stochastic processes

R.H. Stockbridge*

Department of Mathematical Sciences, University of Wisconsin Milwaukee, Milwaukee, WI 53201-0413, USA

Abstract

The analysis of partially observed stochastic control problems often replaces the unknown state process with its conditional distribution given the observations. This technique rewrites the dynamics in terms of knowable processes whose costs coincide with the original processes. This paper considers stochastic processes having singular behavior and presents an approach which separates the determination of the optimal control from the task of estimating the conditional distribution of the unknown process. It involves using a martingale problem characterization for the dynamics followed by a further characterization using occupation measures. This final characterization forms the basis for an equivalent linear programming formulation of the problem over the space of occupation measures.

Keywords: Partially observed stochastic control; Singular stochastic control; Linear programming separation principle; Long-term average criterion

1. Introduction

Consider a prototypical example in partially observed stochastic control. The state process X and observation process Y are specified by

$$\begin{aligned} \mathrm{d}X_t &= f(X_t, u_t)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t, & X_0 &= x_0, \\ \mathrm{d}Y_t &= h(X_t)\,\mathrm{d}t + \sigma\,\mathrm{d}B_t, & Y_0 &= 0, \end{aligned}$$

* Tel.: +414 229 3528; fax: +414 229 4907.
E-mail address: stockbri@uwm.edu.

0362-546X/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.na.2005.02.051

in which W and B are independent, standard Brownian motion processes and u is a process which "controls" the signal. The goal is to select the control process so as to minimize some cost criterion. This paper considers a long-term average cost

$$J(u) = \limsup_{T\to\infty} T^{-1} E\left[\int_0^T c(X_t, u_t)\,\mathrm{d}t\right]$$

in which the function c gives the cost rate associated with the state of the process and the control selected. *The important restriction on the problem is that the control u_t at time t must be based solely on the information available about the observation process Y up to time t.* This information is denoted by the filtration $\mathscr{G} = \{\mathscr{G}_t\}$ in which $\mathscr{G}_t = \sigma(Y_s : 0 \leqslant s \leqslant t)$.

The standard approach is to convert this problem from one involving partial observations to a completely observed control problem by replacing the unobservable process X with the conditional distribution $\pi = \{\pi_t : t \geqslant 0\}$ of X given the observations $\mathscr{G}$. (If X is a process in finite dimensions, this has the effect of reformulating the problem in infinite dimensions since π_t is typically infinite-dimensional.) A desirable property for the new control problem is one in which the task of optimization can be *separated* from that of determining the conditional distribution process π.

This paper addresses a more general model in which the state and observation process may have singular (with respect to time) behavior. By adopting a linear programming approach the estimation and optimization tasks are separated. This separation enables the decision maker to solve "off-line" for the optimal control as a function in feedback form of Y and π, thus leaving the determination of the conditional probabilities π to be performed in real-time.

The equivalence of the linear programming approach to general stochastic control formulations has been established by Bhatt and Borkar [1], Kurtz and Stockbridge [8], and Stockbridge [11]. A similar equivalence between linear programming and optimal stopping has been established by Cho and Stockbridge [2]. Central to the proofs of equivalence is the existence of stochastic processes corresponding to measures satisfying an adjoint relation. Such an existence result for singular stochastic processes is given by Kurtz and Stockbridge [9]. Manne [10] originally observed that discrete-time, finite-state ergodic stochastic control problems could be reformulated as linear programs and this approach has also been extended to Markov decision processes (e.g. [4–6]). The key idea involves rewriting the optimization problem in terms of occupation measures; an idea that originated with Young [12] for problems in the calculus of variations. Also related to this paper is the filtered martingale problem for the conditional distributions π of X given $\mathscr{G}$ studied by Kurtz and Ocone [7].

2. Formulation of the model

2.1. Dynamics

We use the formulation of the singular stochastic processes of Kurtz and Stockbridge [9] and refer the reader to that article for motivation and examples.